# A DEFINITION OF DIMENSIONALITY AND DISTANCE FOR GRAPHS†

29

LOUIS GUTTMAN*
HEBREW UNIVERSITY

## INTRODUCTION

Consider an arbitrary set A and the Cartesian space AA. Let **G** be a symmetric subset of AA, that is, a symmetric binary relation. The characteristic function of **G** can be defined by the real symmetric matrix S of typical element $s_{ab}$, where

$$(1) \quad s_{ab} = \begin{cases} 1 \text{ if } ab \in \mathbf{G} \\ 0 \text{ if } ab \notin \mathbf{G} \end{cases} \quad (a,b \in A)$$

and

$$(2) \quad s_{ab} = s_{ba} \quad (a,b \in A).$$

For our purposes, it is convenient to consider any element $a$ to be in **G** with itself (*i.e.*, $aa \in \mathbf{G}$), so that

$$(3) \quad s_{aa} = 1 \quad (a \in A).$$

In effect, for each element $a$, relation **G** partitions A into two subsets: those elements that relate with $a$ and those that do not. Let $\mathbf{G}_a$ and $\mathbf{H}_a$ be the two respective subsets for $a$:

$$(4) \quad b \in \begin{cases} \mathbf{G}_a \text{ if } ab \in \mathbf{G} \\ \mathbf{H}_a \text{ if } ab \notin \mathbf{G} \end{cases} \quad (a,b \in A).$$

---

† The present paper was originally written in 1965 while Professor Guttman was on sabbatical at the University of Michigan.

* I am indebted to Professor Frank Harary for the stimulation to think about this problem, in a conversation which introduced me to graph theory and his own definition of dimensionality: Erdös, P., Harary, F., & Tutte, W. T. *The dimension of a graph.* (undated stencilled paper).

Then, of course,

(5) $\mathbf{G}_a \cup \mathbf{H}_a = A, \quad \mathbf{G}_a \cap \mathbf{H}_a = 0 \quad (a \in A).$

Two propositions immediately follow from (4) and the preceding conditions:

**PROPOSITION 1:** $a \in \mathbf{G}_a \quad (a \in A)$

**PROPOSITION 2:** If $b \in \mathbf{G}_a$, then $a \in \mathbf{G}_b \quad (a,b \in A).$

Our definition of distance and dimensionality is intended to provide a systematic geometric method for distinguishing between $\mathbf{G}_a$ and $\mathbf{H}_a$ for all $a$ simultaneously. In so doing, it provides a basis for studying the structure of any graph. Let D be a symmetric matrix of the order of S, with real numbers $d_{ab}$ as elements, satisfying the following distance conditions:

(6) $d_{aa} = 0, \quad d_{ab} \geq 0, \quad d_{ab} = d_{ba} \quad (a,b \in A)$

(7) $d_{ab} < d_{ac}$ whenever $b \in \mathbf{G}_a$ but $c \in \mathbf{H}_a \quad (a,b,c \in A).$

We are now in a position to offer the:

*Definition of Dimensionality and Distance.* Let $E_r$ be a Euclidean space of $r$ dimensions for which it is possible to construct a D satisfying (6) and (7), and let $m$ be the smallest possible value of $r$ for all such spaces. Then $m$ will be called the *dimensionality* of G. The elements of the D for an $E_m$ will be called *distances* among the points of A.

THE EXISTENCE THEOREM

In order that the above definition not be vacuous, it must be shown that at least one $E_r$ exists to satisfy (7). This we shall now do, and establish

**THEOREM 1.** If A has $n$ elements, then there always exists an $E_m$ for G, and $m \leq n-1$.

For the proof, let $-\lambda$ be the smallest latent root of the matrix S defined by (1) and (3), and let I denote the unit matrix. Define S* by

(8) $S^* = S + \lambda I.$

Then $S^*$ is a *singular Gramian* matrix; it is not only symmetric but all its latent roots are nonnegative, and at least one root vanishes. These are consequences of the fact that for each latent root $\lambda'$ of S there is a root $\lambda' + \lambda$ for $S^*$, so by the definition of $\lambda$ it must be that $\lambda' + \lambda \geq 0$, the equality holding at least once when $\lambda' = -\lambda$. Consequently, if $r$ is the rank of $S^*$, then $r \leq n-1$, and there exists an $n$ x $r$ matrix X such that

(9) $S^* = XX'$,

where $X'$ is the transpose of X.

Let us pause in the proof for a moment to comment on the sign of the smallest latent root of S. We have defined this to be the negative of $\lambda$, in order that $\lambda$ itself will in general be a positive number; for the smallest root itself is generally negative, as asserted in

**Lemma 1.** If at least one off-diagonal element of S is unity (*i. e.*, $S \neq I$), then $\lambda \geq 0$ (*i. e.*, $-\lambda \leq 0$).

To see this, consider a second-order principal minor of $S^*$ involving a unit off-diagonal element; since $S^*$ is Gramian, such a determinant must be nonnegative:

$$(10) \qquad \begin{vmatrix} 1+\lambda & 1 \\ 1 & 1+\lambda \end{vmatrix} \geq 0.$$

Expanding the determinant establishes the inequality of Lemma 1.

An interesting variation on Lemma 1, which will be used later, is:

**Lemma 2.** A necessary and sufficient condition that $S = I$ is that $\lambda < 0$ (in which case $\lambda = -1$).

The necessity follows from the fact that all roots of I, including $-\lambda$,

equal unity. The sufficiency comes from Lemma 1.

Returning to the proof of Theorem 1, let $x_{ai}$ be the typical element of an X satisfying (9) ($a \in A$; $i=1, 2,\ldots,r$), and let $d_{ab}$ be defined by

$$(11) \qquad d_{ab} = \sqrt{[\Sigma_i (x_{ai} - x_{bi})^2]} \qquad (a,b \in A).$$

Now, from (3), (8), and (9),

$$(12) \qquad \Sigma_i x_{ai}^2 = \lambda + 1 \qquad (a \in A)$$

and

$$(13) \qquad \Sigma_i x_{ai} x_{bi} = s_{ab} \qquad (a \neq b).$$

Expanding the summation in the right of (11), and using (12) and (13), yield

$$(14) \qquad d_{ab} = \sqrt{[2(\lambda + 1 - s_{ab})]} \qquad (a \neq b).$$

Clearly, (14) satisfies (7), for

$$(15) \qquad d_{ab} = \begin{cases} \sqrt{2\lambda} & s_{ab} = 1 \\ \sqrt{2(\lambda+1)} & s_{ab} = 0 \end{cases} \qquad (a \neq b).$$

Definition (11), of course, satisfies (6) as well, and hence provides an explicit $E_r$ into which to map G. Now that at least one $E_r$ is known to exist, clearly there must be a minimal value $m$ for $r$. The present $r$ cannot exceed $n-1$, so neither can $m$; Theorem 1 is fully established.

The construction of (14) yields only two different distances in this particular $E_r$ – apart from the vanishing self-distances – namely those

stated in (15). Smaller dimensionality than $r$ can generally be obtained by constructing more varied distances and yet continuing to satisfy (7).

Although there exists a unique value for minimal $m$ for the dimensionality of G, the distance function for obtaining this is in general not unique. However, certain restraints on D will in general exist (except for disjoint graphs discussed below), for (7) is a special case of what Coombs [1976] calls an "ordered metric". The distances are determined in general up to a set of intervals. Adding a further requirement, such as that of "equal radii" discussed below, may often yield a unique metric.

Of particular interest is the fact that effective computing procedures are available for determining an $E_m$, including programs for high-speed electronic computers [Guttman, 1968; Lingoes, 1973]. These even allow for the presence of "error" in the matrix S, and can provide "approximately" a minimal $E_m$ when desired for empirical data, wherein (7) holds for "almost all" triplets <abc>.

In a fixed $E_m$, each $G_a$ of (4) has associated with it a maximal distance which we shall denote by $\rho_a$, and which we shall call the *radius* of $G_a$:

$$(16) \qquad \rho_a = \max_{b \epsilon G_a} d_{ab} \qquad (a \epsilon A).$$

To know whether or not an element $b$ is in the relation G with $a$, it suffices to know whether or not $d_{ab} \leq \rho_a$. In $E_m$, drawing an $m$-dimensional sphere of radius $\rho_a$, with point $a$ as center, partitions $E_m$ into two subspaces; all points of A outside the sphere are in $H_a$ and the remaining ones are in $G_a$. These remarks may be summarized as

**PROPOSITION 3.** A necessary and sufficient condition that $b \epsilon G_a$ is that $d_{ab} \leq \rho_a$ $(a,b \epsilon A)$.

## EQUIVALENCE CLASSES

Reduction in dimensionality below $n$-1 can always be effected if at least two elements of A are equivalent in a sense now to be defined. This sense has considerable theoretical as well as practical importance.

**Definition of G-equivalence:** Two elements $a$ and $b$ of A will be called G-equivalent if and only if $G_a = G_b$.

From this definition obviously follows

**PROPOSITION 4.** G-equivalence is an equivalence relation.

Hence, all elements of A that are G-equivalent to a given element $a$ constitute an equivalence class, which will be called a *G-class*. The notion of G-equivalence defines a partition of A into mutually exclusive and exhaustive G-classes. Within each G-class, the relational structure is completely known: each element is in the relation with every other element. Internal structures of G-classes differ only on account of the *number* of elements involved. The number of elements in a G-class will be called the *frequency* of the class.

More interesting, and difficult to study, is the structure *between* G-classes. For this purpose, it turns out highly useful that our definition of dimensionality enables us to treat an entire G-class as but a single point in an $E_m$, according to

**THEOREM 2.** If a, b, c, ... are all members of the same G-class, then there exists an $E_m$ in which $d_{ab} = d_{ac} = d_{bc} = \ldots = 0$, and within all G-classes of G simultaneously.

The proof consists of considering a matrix X of order $n$ x $m$ which defines an $E_m$ for G. If $x_{ai}$ (i=1,2,...,$m$) are coordinates of element $a$ in this space, and if $b$ is G-equivalent to $a$, then regardless of what the $x_{bi}$ coordinates may be, they can be replaced by equally good coordinates $x^*_{bi}$ by setting $x^*_{bi} = x_{ai}$ (i=1,2,...,$m$). This leaves all elements of D unchanged save for making $d_{ab} = 0$, so (7) remains satisfied. Making two rows identical cannot increase the rank of X. Hence, since $m$ is already minimal, the new X must also have rank $m$. Equating coordinates for all elements within each G-class therefore establishes Theorem 2. A further consequence is

**THEOREM 3.** If $n^*$ is the number of G-classes in G, then $m \leq n^* - 1$.

For we now need consider only one element from each G-class and use Theorem 1 on this subset of A.

From now on we shall analyze the structure of G by considering each G-class to be a point. With each point is associated a frequency, but the latter will not enter into the analysis of the present paper.

## ISOLATED CLASSES AND DISJOINT GRAPHS

A most unruly type of G-class is one for which no element is in the relation with any element outside the class. Such a class will be called *isolated* or an *isolate*. All other G-classes will be called *non-isolates*. An isolate is characterized by having a radius of zero when Theorem 2 is

exploited, while non-isolates must have spheres of positive radius. The principal submatrix in S defined by the rows and columns of the elements of an isolate consists solidly of entries of unity, and all entries vanish outside this submatrix in these rows and columns.

An isolate is unruly in the sense that there is little restriction on where to locate it in $E_m$, and indeed it does not add anything to the dimensionality of G, as asserted in

**THEOREM 4.** If G has at least two non-isolates, then the dimensionality of G is the dimensionality of its set of non-isolates. If all G-classes are isolates, then the dimensionality of G is one or zero.

To establish the first part of the theorem, consider all non-isolates to be points in some space minimal for them. Then the isolates can be placed at any distinct points in this space outside the spheres of the non-isolates and satisfy (7), leaving the dimensionality unchanged.

If all G-classes are isolates, G will be called *disjoint*. The second part of Theorem 4 asserts that the dimensionality of a disjoint G cannot exceed unity. If A itself is a G-class (all elements of A relate with each other, or G = AA), then it is the only G-class of itself and hence an isolate. Then $m = 0$, for A is but a single point with radius zero. If a disjoint G has two or more isolates, these can always be plotted on a straight line, and in an infinite number of ways. All that is required is that distances be positive between any two points, no matter in what order they are placed, and (7) will be satisfied since $d_{ab} = 0$ whenever $b \in G_a$. From a structural point of view, such a mapping is trivial since it adds no information about the relationship between isolates.

In general, a good strategy by which to analyze graphs is first to determine the isolates and remove them, and then study the structure of what is left.

It is interesting that it is possible to determine whether or not G is disjoint without first determining the G-classes; it suffices to determine the sign of the smallest latent root of S, according to

**THEOREM 5.** A necessary and sufficient condition for G to be disjoint is that S be Gramian. In such a case, the smallest latent root of S is either unity (with $S=I$) or zero (with $S \neq I$).

For the necessity, suppose the rows and columns of S are rearranged to keep those of G-classes together. Then S will have non-overlapping

principal submatrices which are filled uniformly with entries of unity, while all other entries in S vanish. Such an S is obviously Gramian, or $-\lambda \geq 0$. The simplest case is where each G-class has a frequency of one, so that $S = I$ (and $-\lambda = 1$). If $S \neq I$, then S must be singular as well as Gramian, for at least two rows are identical, making $\lambda = 0$.

The sufficiency is given by Lemma 2 when $\lambda < 0$, for then $S = I$. When $\lambda$ vanishes, set $\lambda = 0$ in (15); this makes $d_{ab} = 0$ for any element $b$ in the relation with $a$. Therefore $b$ must have the same coordinates as $a$ in X in (9). But now $S^* = S$ in (8) so that $S = XX'$. Hence $b$ must have the entries in S that $a$ has. More generally, any pair of elements in the relation with each other must be G-equivalent to each other, which implies further that no element in one G-class can be in the relation with an element of another G-class: all G-classes must be isolates. This completes the proof of sufficiency for Theorem 5.

## EXAMPLES OF EUCLIDEAN SPACES FOR GRAPHS

To illustrate the appropriateness of our dimensionality definition for studying graphs, let us consider some of the simplest, but nevertheless important, examples. We shall ignore isolates and frequencies here. First, consider some graphs plottable on a straight line. Two examples of characteristic functions for such are defined by $S_1$ and $S_2$, where

$$S_1 = \begin{bmatrix} 1&1&0&0&0&0 \\ 1&1&1&0&0&0 \\ 0&1&1&1&0&0 \\ 0&0&1&1&1&0 \\ 0&0&0&1&1&1 \\ 0&0&0&0&1&1 \end{bmatrix} \begin{matrix} a \\ b \\ c \\ d \\ e \\ f \end{matrix} \qquad S_2 = \begin{bmatrix} 1&1&1&0&0&0 \\ 1&1&1&1&0&0 \\ 1&1&1&1&1&0 \\ 0&1&1&1&1&0 \\ 0&0&1&1&1&1 \\ 0&0&0&0&1&1 \end{bmatrix} \begin{matrix} a \\ b \\ c \\ d \\ e \\ f \end{matrix}$$

$$\begin{matrix} a & b & c & d & e & f \end{matrix} \qquad \begin{matrix} a & b & c & d & e & f \end{matrix}$$

Beneath each column of $S_1$ and $S_2$ is indicated the name of an element of A to which it belongs. The corresponding graphs $G_1$ and $G_2$ can be portrayedas in Figure 1 below.

*Figure 1*

Graphs for $S_1$ and $S_2$

In each case, there is a unique linear order (except for direction) for the six elements permitted by $S_1$ and $S_2$. In each case, it was also possible to determine a spacing between ranks that would make all $\rho_a$ mutually equal. For $G_1$, the elements are equidistant, and all radii equal $d_{ab}$. Thus, $G_a = \{a,b\}$, $G_b = \{a,b,c\}$, $G_c = \{b,c,d\}$, *etc.*. For $G_2$, by making $d_{ef}$ twice all the other distances, it was again possible to achieve equal radii for all spheres, this time with $\rho = 2d_{ab} = d_{ac}$. Thus, $G_a = \{a,b,c\}$, $G_b = \{a,b,c,d\}$, $G_c = \{a,b,c,d,e\}$, *etc.*, reproducing $S_2$.

Once an $E_m$ is established, introduction of additional constraints (such as equal radii) may often be feasible to determine unique relative distances, as these two samples show. The computer program can be made to seek an $E_m$ with as equal radii as possible. (Alternatively, the program can determine the smallest space in which all radii are equal. This is an alternative approach to the definition of dimensionality which has many virtues.)

Two further examples are of simple, but again important, types of two dimensional graphs, defined by $S_3$ and $S_4$:

$$S_3 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{matrix} \mathbf{a} \\ \mathbf{b} \\ \mathbf{c} \\ \mathbf{d} \\ \mathbf{e} \\ \mathbf{f} \end{matrix}$$

$$\mathbf{a\ b\ c\ d\ e\ f}$$

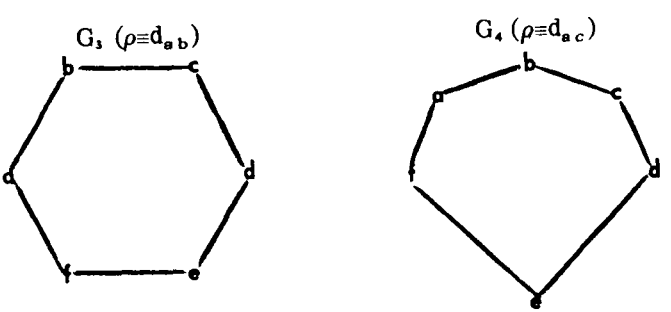
$$S_4 = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{matrix} \mathbf{a} \\ \mathbf{b} \\ \mathbf{c} \\ \mathbf{d} \\ \mathbf{e} \\ \mathbf{f} \end{matrix}$$

$$\mathbf{a\ b\ c\ d\ e\ f}$$

$S_3$ differs from $S_1$ only by the entry in the upper right-hand (and lower left-hand) corner, but this suffices to require a two-dimensional space such as given in Figure 2 for $G_3$.

***Figure 2***

Graphs for $S_3$ and $S_4$



$S_4$ can also be portrayed by a circular arrangement. Again, a uniform radius of $\rho = d_{ab}$ is possible for $G_3$, yielding a unique regular hexagon. While a uniform radius of $\rho = d_{ac}$ is possible for $G_4$, this can be attained in several ways, for $a$ and $c$ can be shifted in the diagram, keeping $d_{ac}$ constant, and yet conform to $S_4$; only the relative positions of $b$, $d$, $e$, and $f$ are fixed by the restraint to a constant radius in this case.

The above illustrations show a basic feature that will hold in the Euclidean representation of any graph. Any points in a straight line in $E_m$ will define a principal submatrix of S of the general nature of $S_1$ and $S_2$: entries adjacent to the main diagonal will be solid blocks of unity, and all others vanish. This will actually hold for any points on an arc that does not bend back on itself. Similarly, any points in a circle, or approximate circle, will define a principal submatrix of S of the general nature of $S_3$ and $S_4$. One is spared having to draw lines between points that relate with each other and trace their paths of interconnections as is done traditionally in graph theory to reach such configurational conclusions (see: chapter on manifolds by Lingoes and Borg). The particular algorithm used to obtain these representations is of some moment here, since there are two sets of ties and weaker constraints and/or $m$ too small can seriously degrade the configuration so that (7) is satisfied by $\leq$ trivially (see: Lingoes & Roskam (this book) on algorithm comparisons).

## ASYMMETRIC RELATIONS

Extension of our definition to the asymmetric case can be done in a number of ways. Because the concept of a relation is so general it is to be expected that different contexts may find different extensions more useful. One kind of extension is as follows:

**Extension 1.** If R is any subset of AA (that is, R is a relation), let $R_a$ be the subset of A of all elements b such that $ab \in R$. Again set $a \in R_a$, so that each element is in the relation with itself. Replace (7) by

(7a) $d_{ab} < d_{ac}$ whenever $b \in R_a$ and/or $a \in R_b$ but $c \notin R_a$ and $a \notin R_c$.

This makes points farther apart when they do not relate to each other at all than when there is a relation in at least one direction. If desired, direction can be indicated in the representation by arrowheads, as in the conventional portrayal of digraphs. A related possibility is:

**Extension 2.** Replace (7) by

(7b) $d_{ab} < d_{ac} < d_{ae}$ whenever the following conditions hold:

(1) $a \in R_b$ and $b \in R_a$
(2) either $a \in R_c$ or $c \in R_a$ but not both
(3) $a \notin R_e$ and $e \notin R_a$.

Condition (7b) enables us to associate with each *a* element *two* radii, one defining a sphere enclosing all points which relate symmetrically with *a*, and a larger one enclosing as well all points that relate with *a* in at least one direction. The difference between these two spheres gives a zone containing the one-directional relations only.

Extension 2 brings out how the whole present treatment of relations is but a special case of the treatment of ordered distances [Guttman, 1968]. Matrix S starts with but two ranks for non-self distances. Condition (7b) provides three initial ranks. The general problem is of an arbitrary number of ranks, and the nonmetric Euclidean technique "unties" tied ranks in an optimal manner for the dimensional problem.

## REFERENCES

Coombs, C. H. *A Theory of Data.* Ann Arbor: Mathesis Press, 1976.

Guttman, L. A general nonmetric technique for finding the smallest Euclidean space for a configuration of points. *Psychometrika*, 1968, *33*, 469-506.

Lingoes, J. C. *The Guttman-Lingoes Nonmetric Program Series.* Ann Arbor: Mathesis Press, 1973.